

WHAT IS CLAIMED IS:

1. A program product for speech recognition, comprising machine-readable program code for causing, when executed, a machine to perform the following method:

detecting at least one speech element in an utterance of acoustic data which meets a criteria for potentially being reduced;

presenting at least a portion of the utterance of acoustic data to a user with a prompt to determine if the speech element is reduced; and

if verification data is received from the user that the speech element in the utterance is reduced, training an acoustic model for a reduced form of the speech element using data related to the utterance.

2. The program product as defined in claim 1, wherein the criteria is that a portion of the acoustic data of the utterance matches data in a dictionary of reduced speech elements.

3. The program product as defined in claim 2, wherein the matched data in the dictionary comprises acoustic data for a reduced speech element.

4. The program product as defined in claim 2, wherein the matched data in the dictionary comprises word sequence context data.

5. The program product as defined in claim 2, wherein the dictionary is external to a speech recognition model in a base speech recognition process.

6. The program product as defined in claim 1, wherein said acoustic model for said reduced form is a discriminative model distinguishing said reduced form from the corresponding unreduced form.

7. The program product as defined in claim 1, wherein the criterion is that the speech element has a substantially lower amplitude than in an unreduced model for that speech element.

8. The program product as defined in claim 1, wherein the criterion is that the speech element has a substantially shorter duration than an average duration for the speech element.

9. The program product as defined in claim 1, wherein the criterion is that the speech element has acoustic characteristics that are less extreme than an unreduced model of the speech element.

10. The program product as defined in claim 1, wherein the criterion is that the speech element has acoustic characteristics that are more like an average speech sound than is an unreduced model for the speech element.

11. The program product as defined in claim 1, wherein the speech element is a vocalic and the criterion is that the speech element has acoustic characteristics more similar to a uniform acoustic tube than does an unreduced model for the speech element.

12. The program product as defined in claim 1, wherein the criterion is that the speech element has acoustic characteristics associated with an incomplete articulatory gesture.

13. The program product as defined in claim 1, wherein the unreduced speech element is formed by closure between the tongue and the roof of the mouth and the criterion is that the speech element has acoustic characteristics associated with only an incomplete or brief contact of the tongue with the roof of the mouth.

14. A program product for speech recognition, comprising machine-readable program code for causing, when executed, a machine to perform the following method:

receiving a training utterance of acoustic data of a word sequence;  
detecting if the training utterance of acoustic data has at least one speech element which meets a criterion for potentially being reduced;

presenting the utterance of acoustic data to a user with a prompt to determine if the speech element is reduced; and

if verification data is received from the user that the speech element in the utterance is reduced, then associating a reduced designation with the speech element in the training utterance designated as reduced.

15. The program product as defined in claim 14, wherein the criterion is that the speech element has a substantially lower amplitude than in an unreduced model for that speech element.

16. The program product as defined in claim 14, wherein the criterion is that the speech element has a substantially shorter duration than an average duration for speech element.

17. The program product as defined in claim 14, wherein the criterion is that the speech element has acoustic characteristics that are less extreme than an unreduced model for the speech element.

18. The program product as defined in claim 14, wherein the criterion is that the speech element has acoustic characteristics that are more like an average speech sound than is an unreduced model for the speech element.

19. The program product as defined in claim 14, wherein the speech element is a vocalic and the criterion is that the speech element has acoustic characteristics more similar to a uniform acoustic tube than does an unreduced model for the speech element.

20. The program product as defined in claim 14, wherein the criterion is that the speech element has acoustic characteristics associated with an incomplete articulatory gesture.

21. The program product as defined in claim 14, wherein the unreduced speech element is formed by closure between the tongue and the roof of the mouth and the criterion is that the speech element has acoustic characteristics associated with only an incomplete or brief contact of the tongue with the roof of the mouth.

22. A speech recognition process, comprising:

detecting at least one speech element in an utterance of acoustic data which meets a criteria for potentially being reduced;

presenting at least a portion of the utterance of acoustic data to a user with a prompt to determine if the speech element is reduced; and

if verification data is received from the user that the speech element in the utterance is reduced, training a discrimination model using data related to the utterance.

23. The method as defined in claim 22, wherein the criteria is that a portion of the acoustic data of the utterance matches data in a dictionary of reduced speech elements.

24. The method as defined in claim 23, wherein the matched data in the dictionary comprises acoustic data for a reduced speech element.

25. The method as defined in claim 23, wherein the matched data in the dictionary comprises word sequence context data.

26. The method as defined in claim 22, wherein the dictionary is external to a speech recognition model in a base speech recognition process.

27. The method as defined in claim 22, wherein said acoustic model for said reduced form is a discriminative model distinguishing said reduced form from the corresponding unreduced form.

28. The method as defined in claim 22, wherein the criterion is that the speech element has a substantially lower amplitude than in an unreduced model for that speech element.

29. The method as defined in claim 22, wherein the criterion is that the speech element has a substantially shorter duration than an average duration for the speech element.

30. The method as defined in claim 22, wherein the criterion is that the speech element has acoustic characteristics that are less extreme than an unreduced model of the speech element.

31. The method as defined in claim 22, wherein the criterion is that the speech element has acoustic characteristics that are more like an average speech sound than is an unreduced model for the speech element.

32. The method as defined in claim 22, wherein the speech element is a vocalic and the criterion is that the speech element has acoustic characteristics more similar to a uniform acoustic tube than does an unreduced model for the speech element.

33. The method as defined in claim 22, wherein the criterion is that the speech element has acoustic characteristics associated with an incomplete articulatory gesture.

34. The method as defined in claim 22, wherein the unreduced speech element is formed by closure between the tongue and the roof of the mouth and the criterion is that the speech element has acoustic characteristics associated with only an incomplete or brief contact of the tongue with the roof of the mouth.

35. A speech recognition method, comprising:  
receiving a training utterance of acoustic data of a word sequence;

detecting if the training utterance of acoustic data has at least one speech element which meets a criterion for potentially being reduced;

presenting the utterance of acoustic data to a user with a prompt to determine if the speech element is reduced; and

if verification data is received from the user that the speech element in the utterance is reduced, then associating a reduced designation with the speech element in the training utterance designated as reduced.

36. The method as defined in claim 35, wherein the criterion is that the speech element has a substantially lower amplitude than in an unreduced model for that speech element.

37. The method as defined in claim 35, wherein the criterion is that the speech element has a substantially shorter duration than an average duration for speech element.

38. The method as defined in claim 35, wherein the criterion is that the speech element has acoustic characteristics that are less extreme than an unreduced model for the speech element.

39. The method as defined in claim 35, wherein the criterion is that the speech element has acoustic characteristics that are more like an average speech sound than is an unreduced model for the speech element.

40. The method as defined in claim 35, wherein the speech element is a vocalic and the criterion is that the speech element has acoustic characteristics more similar to a uniform acoustic tube than does an unreduced model for the speech element.

41. The method as defined in claim 35, wherein the criterion is that the speech element has acoustic characteristics associated with an incomplete articulatory gesture.

42. The method as defined in claim 35, wherein the unreduced speech element is formed by closure between the tongue and the roof of the mouth and the criterion is that the speech element has acoustic characteristics associated with only an incomplete or brief contact of the tongue with the roof of the mouth.

43. A speech recognition system, comprising:

a detector for receiving at least one speech element in an utterance of acoustic data which meets a criteria for potentially being reduced;

a presentation device for presenting at least a portion of the utterance of acoustic data to a user with a prompt to determine if the speech element is reduced; and

a computer training device that, if verification data is received from the user that the speech element in the utterance is reduced, trains an acoustic model for a reduced form of the speech element using data related to the utterance.

44. The system as defined in claim 43, wherein the criteria is that a portion of the acoustic data of the utterance matches data in a dictionary of reduced speech elements.

45. The system as defined in claim 44, wherein the matched data in the dictionary comprises acoustic data for a reduced speech element.

46. The system as defined in claim 44, wherein the matched data in the dictionary comprises word sequence context data.

47. The system as defined in claim 44, wherein the dictionary is external to a speech recognition model in a base speech recognition process.

48. The system as defined in claim 43, wherein said acoustic model for said reduced form is a discriminative model distinguishing said reduced form from the corresponding unreduced form.

49. The system as defined in claim 43, wherein the criterion is that the speech element has a substantially lower amplitude than in an unreduced model for that speech element.

50. The system as defined in claim 43, wherein the criterion is that the speech element has a substantially shorter duration than an average duration for the speech element.

51. The system as defined in claim 43, wherein the criterion is that the speech element has acoustic characteristics that are less extreme than an unreduced model of the speech element.

52. The system as defined in claim 43, wherein the criterion is that the speech element has acoustic characteristics that are more like an average speech sound than is an unreduced model for the speech element.

53. The system as defined in claim 43, wherein the speech element is a vocalic and the criterion is that the speech element has acoustic characteristics more similar to a uniform acoustic tube than does an unreduced model for the speech element.

54. The method as defined in claim 43, wherein the criterion is that the speech element has acoustic characteristics associated with an incomplete articulatory gesture.

55. The method as defined in claim 43, wherein the unreduced speech element is formed by closure between the tongue and the roof of the mouth and the criterion is that the speech element has acoustic characteristics associated with only an incomplete or brief contact of the tongue with the roof of the mouth.

56. A speech recognition system, comprising:

a receiver for receiving a training utterance of acoustic data of a word sequence;

a detector for detecting if the training utterance of acoustic data has at least one speech element which meets a criterion for potentially being reduced;

a presentation device for presenting the utterance of acoustic data to a user with a prompt to determine if the speech element is reduced; and

logic for, if verification data is received from the user that the speech element in the utterance is reduced, then associating a reduced designation with the speech element in the training utterance designated as reduced.

57. The system as defined in claim 56, wherein the criterion is that the speech element has a substantially lower amplitude than in an unreduced model for that speech element.

58. The system as defined in claim 56, wherein the criterion is that the speech element has a substantially shorter duration than an average duration for speech element.

59. The system as defined in claim 56, wherein the criterion is that the speech element has acoustic characteristics that are less extreme than an unreduced model for the speech element.

60. The system as defined in claim 56, wherein the criterion is that the speech element has acoustic characteristics that are more like an average speech sound than is an unreduced model for the speech element.

61. The system as defined in claim 56, wherein the speech element is a vocalic and the criterion is that the speech element has acoustic characteristics more similar to a uniform acoustic tube than does an unreduced model for the speech element.

62. The system as defined in claim 56, wherein the criterion is that the speech element has acoustic characteristics associated with an incomplete articulatory gesture.

63. The system as defined in claim 56, wherein the unreduced speech element is formed by closure between the tongue and the roof of the mouth and the criterion is that the speech element has acoustic characteristics associated with only an incomplete or brief contact of the tongue with the roof of the mouth.